

新型冠状病毒 SARS-CoV-2 棘突蛋白 免疫原性分析及其多肽疫苗设计研究

许志强¹, 鲁 陈², 胡晓军³, 郭 有², 胡巧丽¹, 朱 颖^{2,4}

(1. 南京医科大学第一附属医院, 江苏 南京 210000; 2. 赣南医学院第一附属医院, 江西 赣州 341000;
3. 赣州市疾病预防控制中心, 江西 赣州 341000; 4. 南昌大学生命科学学院, 江西 南昌 330031)

摘要:目的: 分析 SARS-CoV-2 表面棘突蛋白的结构和 B 细胞表位。方法: 首先利用 ProtParam 对 SARS-CoV-2 棘突蛋白的物理化学性质特征进行分析。随后通过 Clustal 软件, 对 SARS-CoV-2 和 SARS-CoV 棘突蛋白中的功能区域进行分析, 并结合同源模拟, 明确棘突蛋白的空间结构和折叠特征。综合 DNASTar、ABCpred 和 BepiPred 结果, 筛选 2019-CoV 棘突蛋白的线性 B 细胞表位, 同时利用 ElliPro、DiscoTope、SEPPA 对棘突蛋白的构象 B 细胞表位进行综合预测。结果: 通过同源模拟, SARS-CoV-2 的棘突蛋白是同源三聚体结构, 并且与 SARS-CoV 的棘突蛋白相似, 也具有两种受体结合区域构象。通过综合多种免疫信息学工具及进一步筛选, 11 条线性表位 (B9 - B14 和 B27 - B31) 及 5 条构象表位 (CB4 - CB8) 被筛出。结论: 本研究筛选出的 SARS-CoV-2 棘突蛋白的 16 条 B 细胞表位, 包括 11 条线性表位 (B9 - B14 和 B27 - B31) 及 5 条构象表位 (CB4 - CB8), 可作为潜在的新型冠状病毒肺炎免疫原疫苗研发的候选表位。

关键词: SARS-CoV-2; COVID-19; 三维结构模拟; B 细胞表位

中图分类号: R392.11 文献标志码: A

Epitope-based peptide vaccine design and immunogenicity analysis of the spike protein from SARS-CoV-2

XU Zhi-qiang¹, LU Chen², HU Xiao-jun³, GUO You², HU Qiao-li¹, ZHU Ying^{2,4}

(1. The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210000; 2. The First Affiliated Hospital of Gannan Medical University, Ganzhou, Jiangxi 341000; 3. Ganzhou Center for Disease Control and Prevention, Ganzhou, Jiangxi 341000; 4. School of Life Science, Nanchang University, Nanchang, Jiangxi 330031)

Abstract: Objective: To analyse the structure and B-cell epitope of spike protein from SARS-CoV-2 surface. **Methods:** Firstly, the physiochemical properties of the SARS-CoV-2 spike protein were analyzed by the ProtParam. Subsequently, the architecture of the SARS-CoV-2 spike protein was compared with that of SARS-CoV by sequence alignment in Clustal. The tertiary structure of SARS-CoV-2 spike protein was built by homologous modelling. Finally, both linear and conformational epitopes in spike protein were predicted by different immunoinformatic approaches including DNASTar, ABCpred, BepiPred, ElliPro, DiscoTope and SEPPA. **Results:** Through homologous simulation, the spike protein on SARS-CoV-2 was a homo-trimer and had two different conformations of receptor-binding domain which was similar with SARS-CoV. By combining multiple immunoinformatics tools and further screening, eleven linear epitopes (B9 - B14 and B27 - B31) and five conformational epitopes (CB4 - CB8) were screened out. **Conclusions:** Sixteen B cell epitopes of SARS-CoV-2 spike protein, including eleven linear epitopes (B9 - B14 and B27 - B31) and five conformational epitopes (CB4 - CB8), were suggested to be the potential candidate epitopes for the development of novel coronavirus pneumonia immunogenicity vaccine in the study.

Key words: SARS-CoV-2; COVID-19; modelling of 3D structure; B-cell epitopes

基金项目: 赣南医学院 COVID-19 应急项目 (YJ202014)

作者简介: 许志强, 男, 硕士, 药师, 研究方向: 药学。E-mail: xu.zhiqiang@outlook.com

通信作者: 朱颖, 女, 硕士, 副主任医师, 研究方向: 生物化学与分子生物学。E-mail: pink-shadow@163.com

至2020年3月4日,我国被确诊感染新型冠状病毒肺炎患者超过8万人,并导致超过3千名患者死亡,且确诊和死亡人数仍在每日增加(<http://www.nhc.gov.cn/>),对国民生活、经济秩序造成极其严重的影响。目前,国际病毒分类委员会已将这种新型冠状病毒命名为“SARS-CoV-2”,而由此病毒感染的肺炎也被命名为“COVID-19”(Corona Virus Disease 2019)。

冠状病毒是一种具有包膜的单股正链RNA病毒。SARS-CoV-2属于 β -冠状病毒,具有典型的冠状病毒基因组结构。棘突蛋白(spike protein, S)位于SARS-CoV-2病毒包膜表面,可与宿主细胞膜特定受体结合,在病毒侵入和感染细胞中具有至关重要的作用。尽管SARS-CoV-2的S蛋白在多种冠状病毒中表现出高度的变异,但是它与引起严重急性呼吸道综合征(severe acute respiratory syndrome, SARS)的SARS-CoV病毒表面的S蛋白具有较高的同源性。SARS-CoV-2与SARS-CoV的S蛋白在其受体结合域(receptor-binding domain, RBD)具有几乎相同的三维空间结构,表明SARS-CoV-2与SARS-CoV作用的受体相同,即都是通过S蛋白-血管紧张素转化酶2(angiotensin-converting enzyme 2, ACE2)结合途径感染支气管上皮细胞和肺上皮细胞^[1]。截至2020年3月5日,SARS-CoV-2感染病例仍在增加,因此,有效的疫苗研发至关重要。

目前,已有大量的药物可作为SARS-CoV-2感染后的辅助治疗,包括广谱的病毒靶向药物,宿主靶向药物和中药等^[2-3]。然而,市场上还没有特异性针对SARS-CoV-2的药物或疫苗,一些新型药物的研制也处于早期阶段。基于S蛋白在SARS-CoV-2病毒感染中的关键作用,其有望成为治疗药物设计靶点。已发表的针对SARS-CoV S蛋白结构的研究表明,其N末端区域(N terminal domain, NTD),融合肽区域(fusion peptide, F),七肽重复区(heptad region, HR)和中央螺旋区(central helix, CH)是激发体内产生中和抗体(neutralizing antibody)的潜在靶标^[4]。因此,对SARS-CoV-2 S蛋白结构和表位的分析可在一定程度上为免疫原疫苗的研制提供依据。

本研究基于SARS-CoV S蛋白结构和SARS-CoV-2的基因组序列,对其S蛋白的功能域、空间结构进行分析,并通过DNAStar、ABCpred、Bepipred、ElliPro、DiscoTope和SEPPA对SARS-CoV-2的线性和构象B细胞表位进行了综合预测。在抗

击COVID-19的紧迫形势下,本研究有望为新型疫苗的开发和研制节省时间和成本。

1 材料和方法

1.1 序列检索与SARS-CoV-2 S蛋白的物理化学性质分析 从美国国家生物技术信息中心数据库(<https://www.ncbi.nlm.nih.gov/>)检索SARS-CoV-2和SARS-CoV两种冠状病毒S蛋白的氨基酸序列,在GenBank中的检索号分别为QHD43416和P59594。利用ProtParam对SARS-CoV-2 S蛋白的分子量,氨基酸组成,亲水性均值和不稳定指数进行分析。

1.2 SARS-CoV-2与SARS-CoV S蛋白的氨基酸序列比对和体系结构分析 我们通过Clustal序列比对工具将SARS-CoV-2与SARS-CoV的S蛋白进行氨基酸序列比对,明确两种病毒中S蛋白的同源性,并根据已报道的SARS-CoV S蛋白结构域^[4],进一步分析SARS-CoV-2 S蛋白的信号肽区域(signal peptide, SP)、N末端区域(N-terminal domain, NTD)、连接区域(linker or connect region, L or CR)、受体结合区域等结构(receptor-binding domain, RBD)。此外,通过TMHMM 2.0(<http://www.cbs.dtu.dk/services/TMHMM/>)对S蛋白的跨膜区域进行了预测^[5]。

1.3 二级结构和三级结构的构建与分析 根据SARS-CoV-2的氨基酸序列,利用SWISS-MODEL构建了SARS-CoV-2 S蛋白的三级结构^[6]。在SWISS-MODEL服务器(<https://swissmodel.expasy.org/>)中输入SARS-CoV-2 S蛋白的氨基酸序列进行同源模拟,选择具有最大序列一致性、最大序列覆盖率、最高质量得分(Global Model Quality Estimation, GMQE)得分最佳模型进行进一步分析。使用pymol展示最终获得的SARS-CoV-2的S蛋白模型^[7],并分析其中 α -螺旋和 β -折叠结构。

1.4 SARS-CoV-2线性B细胞表位分析 B细胞表位是抗原蛋白中能被B细胞受体或随后由免疫系统产生的抗体所识别的特殊区域。一般来说,B细胞表位可分为两种类型,即线性表位(linear epitope)和构象表位(conformational epitope)。线性表位又称连续性表位,是由抗原的连续性的氨基酸组成。对于SARS-CoV-2的S蛋白的线性表位,我们利用三种不同类型的免疫信息学工具:DNAStar protean、ABCpred、Bepipred,全面分析线性B细胞表位。在DNAStar中,选择S蛋白氨基酸具有良好亲水性、柔韧性、可及性和高抗原性的区域作为初步筛

选结果。随后使用 ABCpred^[8]和 Bepipred 2.0^[9]对 S 蛋白的线性表位进一步分析,分别从中选择得分高于 0.8 和 0.55 结果。综合上述三种工具所得结果,选择由两种工具以上筛选的相同序列作为最终的候选线性 B 细胞表位。

1.5 SARS-CoV-2 构象 B 细胞表位分析 不同于线性 B 细胞表位,构象 B 细胞表位可由分布在三维结构中位置临近的非连续性的氨基酸组成。首先结合 S 蛋白三维结构,使用 ElliPro 筛选最小分数为 0.8 和最大距离为 5 埃的构象表位^[10]。随后,使用 DiscoTope 1.1 和 SEPPA 3.0 区分 S 蛋白氨基酸序列中的表位和非表位残基^[11-12]。综合 ElliPro、DiscoTope 2.0、SEPPA 3.0 的筛选结果,确定最终的构象 B 细胞表位。

2 结果

2.1 SARS-CoV-2 S 蛋白的序列检索和物理化学性质 根据 SARS-CoV-2 S 蛋白的 Genbank 检索号 (QHD43416) 获取其氨基酸序列。通过 ProtParam 分析,SARS-CoV-2 的 S 蛋白包含 1 273 个氨基酸,分子量为 141.18 kDa。其氨基酸中,110 个带负电荷(D + E),103 个带正电荷(R + K)。理论等电点(pI)为 6.24,表明它在天然条件下是带负电荷的蛋白质。其亲水性均值和不稳定指数分别为 33.01 和 -0.079,表明 S 蛋白具有稳定和亲水性特性。

2.2 序列对比和体系结构分析 通过 Clustal 软件对 SARS-CoV-2 和 SARS-CoV S 蛋白的氨基酸序列进行比对,它们的氨基酸一致性为 76%,具有较高的同源性。通过与 SARS-CoV 的序列相应区域比较,在 SARS-CoV-2 的 S 蛋白可分为两个部分:S1 和 S2。富含 β 折叠结构的 S1 亚基主要由 N 末端区域(氨基酸残基 18-306)、受体结合区域(氨基酸残基 331-527)、两个亚区域(氨基酸残基 528-589 和 590-676)组成。S2 亚基则主要含有 α 螺旋结构,包括一个上游螺旋结构(upstream helix, UH)位于 747-783、一个螺旋融合区域(fusion region, F)位于 816-827、七肽重复区(heptated region, HR)位于 912-984,中央 α 螺旋区域(central helix)位于 985-1 034。随后紧接一个“发卡”结构(β -hairpin, β H)位于 1 035-1 071 和一个亚区域位于 1 072-1 112。信号肽区、跨膜区和膜内区则分别位于 1-

17、1 214-1 236 和 1 237-1 273(图 1)。

2.3 SARS-CoV-2 S 蛋白的空间结构和折叠特征

在 SWISS-MODEL 中,与 SARS-CoV-2 S 蛋白氨基酸序列一致性和覆盖率最高的模板为 SARS-CoV S 蛋白(PDB 号:6ACC)。通过模型构建,SARS-CoV-2 的 S 蛋白为同源三聚体结构,其中一个单体的受体结合区存在两种构象,即“横卧”和“立式”(图 2B~C)。S1 亚基含有 35 个 β -折叠和 6 个 α -螺旋结构,而 S2 亚基包含 14 个 α -螺旋和 7 个 β -折叠结构,其中第 39 和 40 个 β -折叠共同形成 β -发夹结构(图 2D)。

2.4 线性 B 细胞表位分析 鉴定抗原中的 B 细胞表位是表位疫苗设计的重要步骤。综合亲水性、柔韧性、表面可及性和抗原性在内的序列特性,通过 DNASTar protean 得到 40 条多肽。ABCpred 和 Bepipred 2.0 分别筛选出 43 条和 35 条符合条件的多肽。综合三种工具的肽段结果,最终确定出 31 个线性的 SARS-CoV-2 S 蛋白表位,其首-尾残基位置分别为:73-81、94-100、110-115、146-160、204-208、251-257、280-285、312-320、354-362、402-406、415-419、439-444、476-483、526-540、552-556、565-579、602-609、634-645、652-661、674-689、689-704、741-749、772-780、807-818、883-889、931-946、1 084-1 091、1 117-1 127、1 157-1 164、1 180-1 186、1 191-1 196,每个表位对应的氨基酸序列和在 S 蛋白空间结构上的位置分别见表 1 和图 3A,其中 7 个线性表位 B1-B7 位于 SARS-CoV-2 S 蛋白的 NTD 区域(氨基酸残基 18-306),5 个线性表位 B9-B13 位于 RBD 区域(氨基酸残基 331-527)。

2.5 构象 B 细胞表位分析 对 SARS-CoV-2 S 蛋白的构象表位分析中,ElliPro 共得出 14 条分值高于 0.8 且最大距离不超过 5 埃的表位。此外,通过 DiscoTope 1.1 和 SEPPA 3.0 分别鉴定出 184 和 235 个表位残基。整合这三种方法,共获得 9 条构象 B 细胞表位:72-75、(146-147)-(149-155)、246-257、(404-409)-(413-414)、443-448、458-463、474-481、498-506、556-571,相应的氨基酸序列见表 2,其中 3 条表位 CB1-CB3 位于 S 蛋白 NTD 区域,5 条表位 C4-C8 位于 RBD 区域(图 3B)。

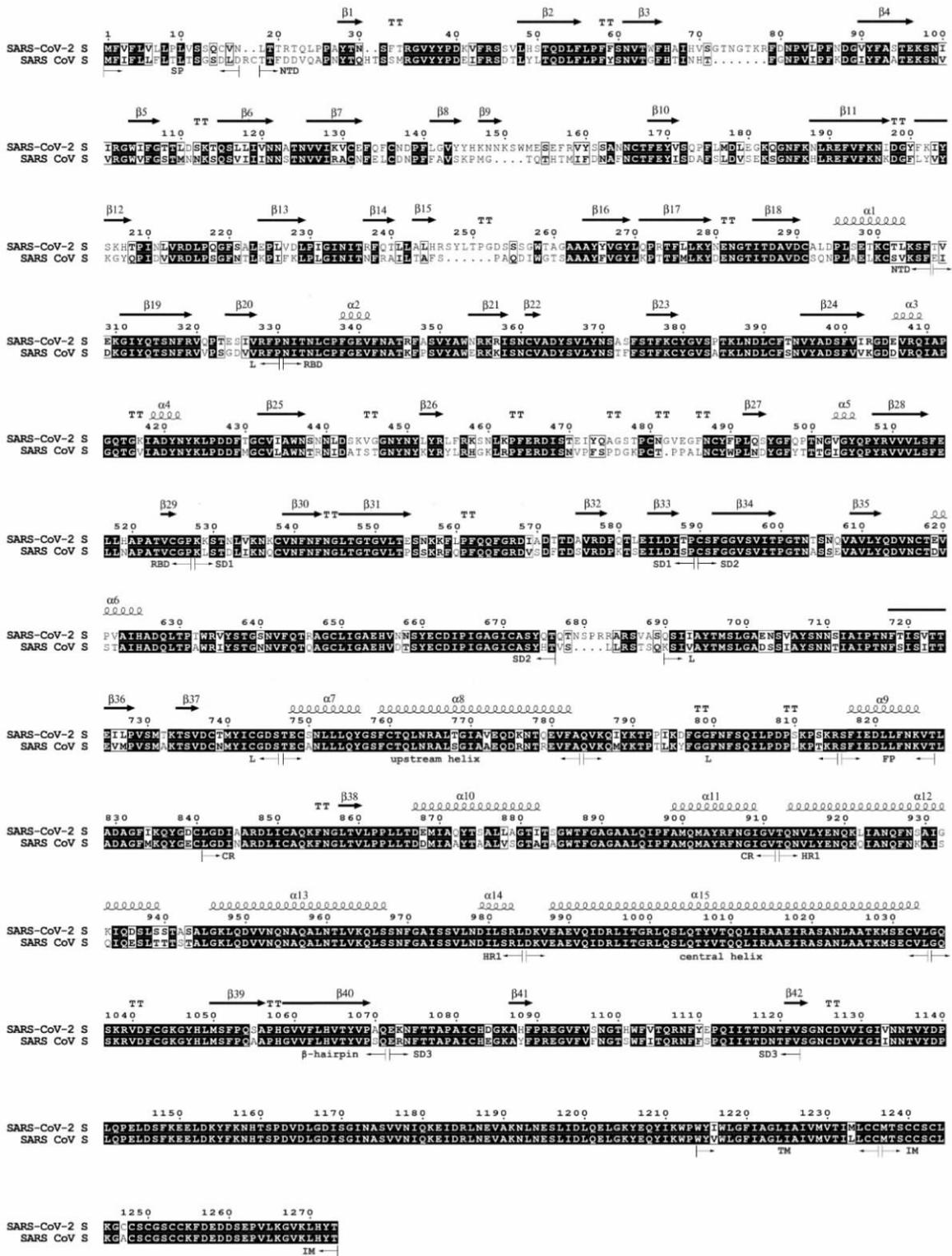


图 1 SARS-CoV-2 与 SARS-CoV S 蛋白的氨基酸序列比对结果

SP: 信号肽;NTD:N 末端区域;L:连接区域;RBD:受体结合区域;SD1:亚区 1;SD2:亚区 2;upstream helix:向上螺旋结构;FP:融合肽;CR:连接区域;HR:七肽重复区;central helix:中央螺旋结构;β-hairpin:β 发卡结构;SD3:亚区 3;α1-α15:α 螺旋结构;β1-β42:β 折叠结构。

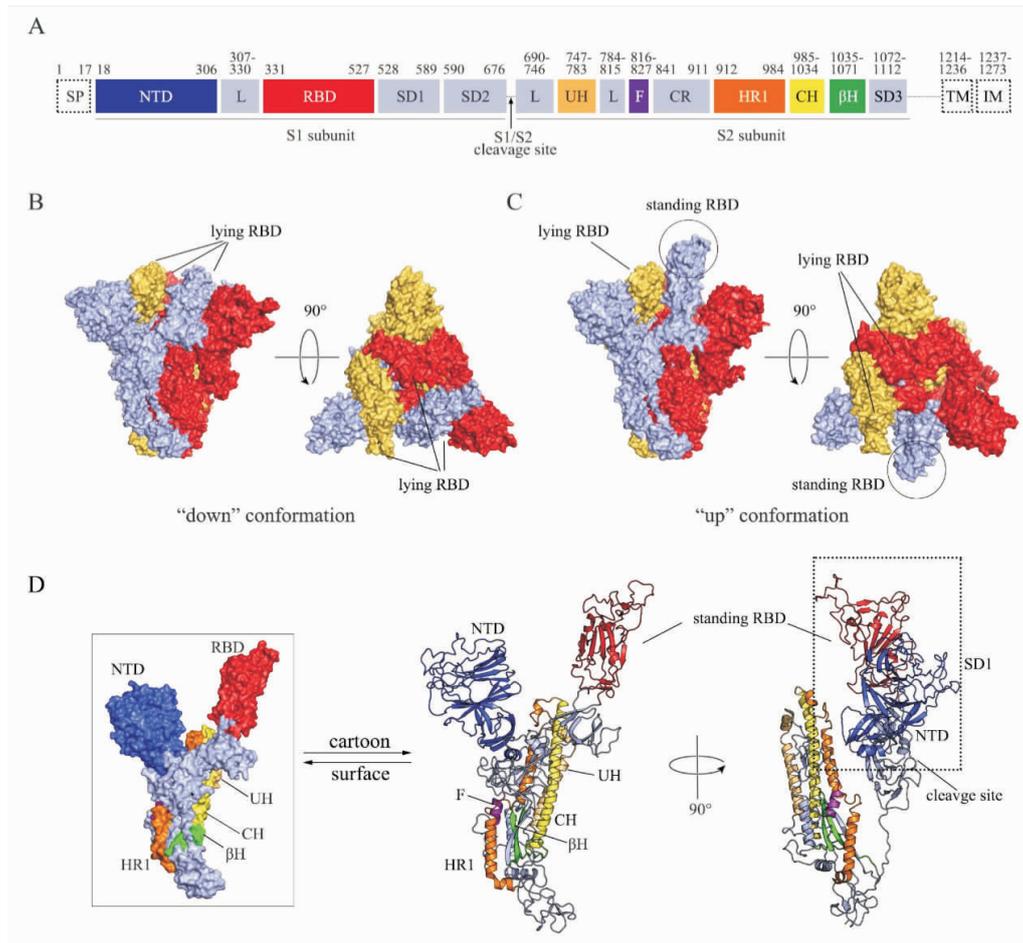


图 2 SARS-CoV-2 S 蛋白的空间结构

A. 不同体系结构的平面分布图示。缩略词释义见图 1。B~C. S 蛋白的三聚体结构和两种构象的蛋白结构。lying RBD:“卧式”受体结合区;standing RBD:“立式”受体结合区。D:S 蛋白不同体系结构的的空间位置。cartoon:以二级结构展示;surface:以表面形式展示。

表 1 线性 B 细胞表位筛选结果

Name	Position(start - end)	Amino acid sequence	Name	Position(start - end)	Amino acid sequence
B1	73 - 81	TNGTKRFDN	B17	602 - 609	TNTSNQVA
B2	94 - 100	STEKSN	B18	634 - 645	RVYSTGSNVFQT
B3	110 - 115	LDSKTQ	B19	652 - 661	GAEHNNSYE
B4	146 - 160	HKNNKSWMESEFRVY	B20	674 - 689	YQTQTNPRRARSVAS
B5	204 - 208	YSKHT	B21	689 - 704	SQSHAYTMSLGAENS
B6	251 - 257	PGDSSSG	B22	741 - 749	YICGDSTEC
B7	280 - 285	NENGTI	B23	772 - 780	VEQDKNTQE
B8	312 - 320	IYQTSNFRV	B24	807 - 818	PDPSKPSKRSFI
B9	354 - 362	NRKRISNCV	B25	883 - 889	TSGWTFG
B10	402 - 406	IRGDE	B26	931 - 946	IGKIQDLSLSTASALG
B11	415 - 419	TGKIA	B27	1 084 - 1 091	DGKAHFPR
B12	439 - 444	NNLSK	B28	1 117 - 1 127	TDNTFVSGNCD
B13	476 - 483	GSTPCNGV	B29	1 157 - 1 164	KNHTSPDV
B14	526 - 540	GPKKSTNLVKNKCVN	B30	1 180 - 1 186	QKEIDRL
B15	552 - 556	LTESN	B31	1 191 - 1 196	KNLNES
B16	565 - 579	FGRDIADTTDAVRDP			

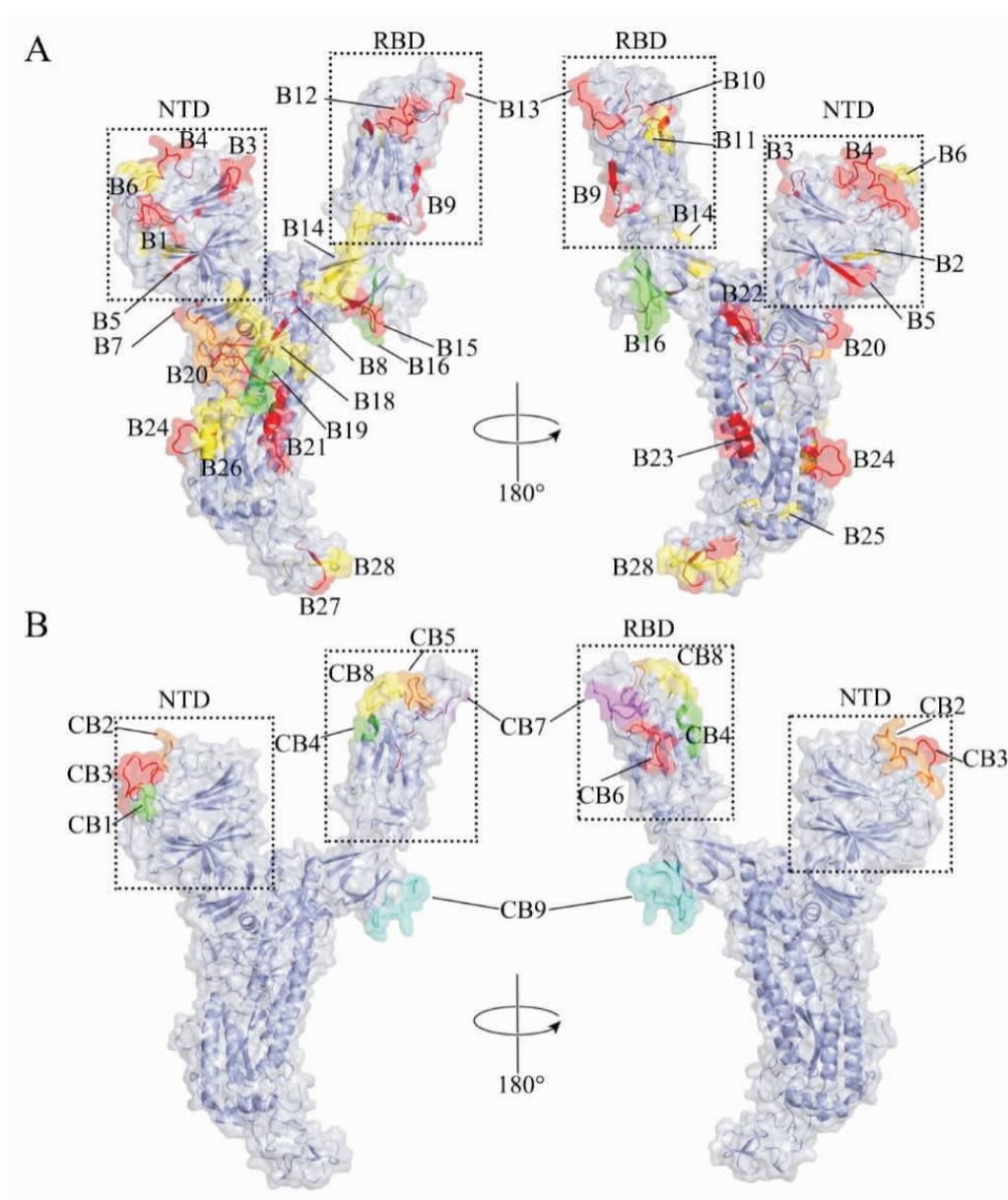


图3 B 细胞表位在 SARS-CoV-2 S 蛋白空间结构上的位置

A. 线性 B 细胞表位在 S 蛋白结构上位置。B1 - B31:31 个不同的 B 细胞表位。B. 构象 B 细胞表位在 S 蛋白结构上的位置。CB1 - CB9:9 个不同的构象 B 表位。

表2 构象 B 细胞表位筛选结果

Name	Position(start - end)	Amino acid sequence
CB1	72 - 75	GTNG
CB2	(146 - 147) - (149 - 155)	(HK)-(NKSWMES)
CB3	246 - 257	RSYLTPGDSSSG
CB4	(404 - 409) - (413 - 414)	(GDEVQR)-(GQ)
CB5	443 - 448	SKVGGN
CB6	458 - 463	KSNLKP
CB7	474 - 481	QAGSTPCN
CB8	498 - 506	QPTNGVGYQ
CB9	556 - 571	NKKFLPFQQFGRDIAD

3 讨论

自 2019 年 12 月至今,新型冠状病毒在国内已引起超过 8 万人感染并导致 3 千名以上患者死亡,对公众健康造成极其严重的危害。目前,2019 新型冠状病毒的基因组信息已经明确,与引起 2003 年严重呼吸综合症的 SARS 冠状病毒有非常高的同源性,并且在感染模式上也具有相似性。由于 S 蛋白在新型冠状病毒传播过程中起着至关重要的作用,因此也是诱导体内中和抗体产生的首要靶点。鉴于疫情防控的严峻形势,迫切地需要明确 SARS-CoV-2 S

蛋白的结构特征和 B 细胞表位,以至于能够快速研制免疫原疫苗和中和性抗体。

我们通过 SARS-CoV 中的 S 蛋白进行氨基酸序列和体系结构的比较,发现虽然 SARS-CoV-2 的 S 蛋白与 SARS-CoV 在总体序列上具有高达 76% 的序列一致性,但在 S 蛋白的 S1 亚基中两者序列差异较大,并存在较多的缺口(gap),序列一致性为 64%。S1 亚基含有 NTD 和与 ACE2 受体结合的 RBD,表明 S1 亚基是 SARS-CoV-2 特异性疫苗研制的重要区域。相较于 S1 亚基,SARS-CoV-2 与 SARS-CoV S 蛋白的 S2 亚基的序列高度相似,氨基酸一致性达到 90%。虽然 Chan 等比较了 SARS-CoV-2 和 SARS-CoV 基因组的相似性并且也针对 S 蛋白进行了序列比对,但是他们主要针对于进化层面进行了分析,对于 SARS-CoV-2 S 蛋白也仅明确了 S1、S2、SP、NTD 和 RBD 区域^[13]。而先前针对 SARS-CoV S 蛋白的研究显示 S2 亚基含有融合肽区域、七肽重复区和中央螺旋结构,均为诱使广谱中和抗体(broadly neutralizing antibodies)产生的重要位点^[4]。因此,本研究还根据 SARS-CoV 的研究结果,同源比较分析得到其他多种功能区所在的位点。

随后,我们通过同源模拟成功构建了 S 蛋白的三维结构模型。与 SARS-CoV 相同,SARS-CoV-2 的 S 蛋白也是一个同源三聚体,并且具有两种构象:其一是“卧式”,三个 RBD 全部都是隐蔽构象,由于位阻效应不易与 ACE2 受体结合;其二是“立式”,三聚体中的一个 RBD 转变为立式构象,是结合至 ACE2 受体的必要条件^[14]。根据 SARS-CoV-2 的 S 蛋白结构模型,我们对其螺旋结构和折叠特征进行了分析。S1 亚基主要为 β -折叠结构而 S2 亚基主要为 α -螺旋结构,这两种结构通常含有较高的化学键能量,不利于表位的形成。Wrapp 等使用冷冻电镜对 SARS-CoV-2 S 蛋白的 27-1146 位氨基酸之间的结构的实际测定结果(PDB ID:6VSB)更是进一步证实了上述结构特征^[15]。我们通过同源模拟所得到的 S 蛋白的结构与电镜实测结构具有较高的一致性,通过均方根偏差计算(RMSD),两种方法所得 S 蛋白的“立式”构象的 RMSD 值为 1.365 Å(750 个原子)，“卧式”构象为 1.494 Å(788 个原子)。两者结构模型之间的差异主要在两个方面:(1)两种模型 S 蛋白 RBD 区域所旋转的角度不同引起,先前对 SARS 病毒 S 蛋白的结构研究中也表明 RBD 区域存在多种旋转角度^[14];(2)虽然冷冻电镜得到的 SARS-CoV-2 S 蛋白的实测结构,但作者表示其中缺

少了一些柔性卷曲结构^[15]。由于 B 细胞表位常处于蛋白的无规则卷曲和 β -转角部分,明确完整的 S 蛋白的二级结构和三维结构对表位疫苗的研制尤为必要。

对于抗原 B 细胞表位的预测一般都基于氨基酸序列的物理化学特性,但仅适于线性 B 表位的分析。然而,构象表位在机体免疫反应过程中同样重要。因此,本研究中不仅对 SARS-CoV-2 的线性 B 细胞表位进行了全面分析,还基于“立式”构象的 S 蛋白结构对构象表位进行了综合预测。利用 DNASTar、ABCpred、BepiPred 三种工具共得到 31 条线性 B 细胞表位,分布于 S 蛋白各处,包括 NTD、RBD、S2 亚基。除 B2 以外,其余 30 条表位均位于 S 蛋白表面和无规则卷曲处。在构象表位的分析中,结合 ElliPro 得到的表位序列、DiscoTope 和 SEPPA 所得阳性氨基酸残基,综合出 9 条构象表位,主要分布于 S1 的 NTD 和 RBD,并且和线性 B 细胞表位具有部分重合,如 B1 与 CB1、B4 与 CB2、B6 与 CB3、B13 与 CB7 等。综合先前已报道对 SARS-CoV S 蛋白的 B 表位研究结果,其 310-535、604-625、1023-1189 均为验证后的表位区域^[16-18],其中 471-503、604-625、1164-1191 在恒河猴中诱导产生的抗体可有效阻止 SARS-CoV 的感染^[17]。这些位点在 SARS-CoV-2 S 蛋白中相应的位置为 323-549、618-639、1041-1207,通过对表位的进一步筛选,位于这些区域的线性表位共有 11 条(B9-B14 和 B27-B31)、构象表位共有 5 条(CB4-CB8),均为潜在的疫苗研发候选表位。在这 16 条表位中,B9-B13、CB4-CB8 均位于 SARS-CoV-2 S 蛋白与人 ACE2 受体结合的 RBD 区域,针对此 10 条表位疫苗产生的抗体在理论上能直接干扰 S 蛋白与人 ACE2 的结合。虽然 BARUAH 等也对 SARS-CoV-2 S 蛋白的 B 细胞表位进行了分析,但是缺少全面性和综合性。在线性表位筛选中没有考虑氨基酸自身的物理化学特性,在构象表位中也仅依靠 ElliPro 单一工具所得结果,最终只获得了 3 条线性和 5 条构象 B 细胞表位^[19]。较少表位结果既丧失了灵敏性也缺失准确性,不利于后续疫苗的研究。本研究不仅综合了多种工具在多种 S 蛋白的特性下对表位进行了综合分析,而且根据先前针对 SARS S 蛋白 B 表位已有的研究和 SARS-CoV-2 的功能区,对所得表位结果进行了进一步的筛选,旨在加速研发针对 SARS-CoV-2 的中和抗体。

综上所述,我们不仅对 SARS-CoV-2 S 蛋白的结

构域进行了分析,还对潜在表位进行了综合筛选,共获得40条B细胞表位。综合SARS-CoV的表位研究结果,我们从中筛选出16条可作为潜在的新型冠状病毒肺炎免疫原疫苗研发的候选表位,且其中的10条位于SARS-CoV-2 S蛋白RBD区,是较为有进一步研究价值的肽段。这些结果不仅能减少疫情防控所需疫苗的研制时间,更能进一步节约其研发成本。

参考文献:

- [1] XU X, CHEN P, WANG J, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission[J]. *Sci China Life Sci*, 2020, 63(3):457-460.
- [2] LI GD, CLERCQ ED. Therapeutic options for the 2019 novel coronavirus (2019-nCoV)[J]. *Nat Rev Drug Discov*, 2020, 19(3):149-150.
- [3] LU H. Drug treatment options for the 2019-new coronavirus (2019-nCoV)[J]. *Bioscience trends*, 2020. doi:10.5582/bst.2020.01020.
- [4] YUAN Y, CAO D, ZHANG Y, et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains[J]. *Nat Commun*, 2017, 8:15092.
- [5] KROGH A, LARSSON B, VON HEIJNE G, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes[J]. *J Mol Biol*, 2001, 305(3):567-580.
- [6] WATERHOUSE A, BERTONI M, BIENERT S, et al. SWISS-MODEL: homology modelling of protein structures and complexes[J]. *Nucleic Acids Res*, 2018, 46(W1):W296-W303.
- [7] DELANO W L. Pymol: An open-source molecular graphics tool[J]. *CCP4 Newsletter on protein crystallography*, 2002, 40(1):82-92.
- [8] SAHA S, RAGHAVA GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network[J]. *Proteins*, 2006, 65(1):40-8.
- [9] JESPERSEN MC, PETERS B, NIELSEN M, et al. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes[J]. *Nucleic Acids Res*, 2017, 45(W1):W24-W29.
- [10] PONOMARENKO J, BUI HH, LI W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes[J]. *BMC Bioinformatics*, 2008, 9:514.
- [11] KRINGELUM JV, LUNDEGAARD C, LUND O, et al. Reliable B cell epitope predictions: impacts of method development and improved benchmarking[J]. *PLoS Comput Biol*, 2012, 8(12):e1002829.
- [12] ZHOU C, CHEN Z, ZHANG L, et al. SEPPA 3.0-enhanced spatial epitope prediction enabling glycoprotein antigens[J]. *Nucleic Acids Res*, 2019, 47(W1):W388-W394.
- [13] CHAN J F W, KOK K H, ZHU Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan[J]. *Emerging Microbes & Infections*, 2020, 9(1):221-236.
- [14] GUI M, SONG W, ZHOU H, et al. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding[J]. *Cell Res*, 2017, 27(1):119-129.
- [15] WRAPP D, WANG N, CORBETT K S, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation[J]. *Science*, 2020.
- [16] DUAN J, YAN X, GUO X, et al. A human SARS-CoV neutralizing antibody against epitope on S2 protein[J]. *Biochem Biophys Res Commun*, 2005, 333(1):186-193.
- [17] CHOU TH, WANG S, SAKHATSKYY PV, et al. Epitope mapping and biological function analysis of antibodies produced by immunization of mice with an inactivated Chinese isolate of severe acute respiratory syndrome-associated coronavirus (SARS-CoV)[J]. *Virology*, 2005, 334(1):134-143.
- [18] WANG Q, ZHANG L, KUWAHARA K, et al. Immunodominant SARS Coronavirus Epitopes in Humans Elicited both Enhancing and Neutralizing Effects on Infection in Non-human Primates[J]. *ACS Infect Dis*, 2016, 2(5):361-376.
- [19] BARUAH V, BOSE S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV[J]. *J Med Virol*, 2020. doi:10.1002/jmv.25698.

(收稿日期:2020-02-17)(责任编辑:敖慧斌)